

Chapter 8

The impact of nominalisations on the reading process: A case-study using the Freiburg Legalese Reading Corpus

Sascha Wolfer

Institute for the German Language, Mannheim

I present a study using eye-tracking-while-reading data from participants reading German jurisdictional texts. I am particularly interested in nominalisations. It can be shown that nominalisations are read significantly longer than other nouns and that this effect is quite strong. Furthermore, the results suggest that nouns are read faster in reformulated texts. In the reformulations, nominalisations were transformed into verbal structures. Reformulations did not lead to increased processing times of verbal constructions but reformulated texts were read faster overall. Where appropriate, results are compared to a previous study of Hansen et al. (2006) using the same texts but other methodology and statistical analysis.

1 Introduction

In linguistics, text corpora are used to analyse rules and usage of language in natural contexts. Most of the time, corpora are annotated with linguistic information. These annotations allow linguists to search for recurring patterns and extract all instances of a specific linguistic structure from the corpus. These linguistic annotations can span several levels of linguistic structures (e.g. parts-of-speech and phrase structure). On the textual level, one might be interested in co-reference chains to investigate which words in the text refer to the same entity in the world. Some annotations can also be numerical in nature. One of the most common measures associated with words is the frequency with which the word occurs in natural language. Most of the time, a corpus itself is the source of this information. If one compiles a corpus of terminological language, however, the frequency of words in everyday language might also be interesting. Here, the source of the frequency information might well be another corpus.



In recent years, some researchers began to enrich corpora with another kind of annotation layer. If I collect or compile texts to create a corpus of natural language, I can take this corpus and show its contents to humans. Of course, these humans should be taken from a population likely to be confronted with the kind of text material I included in the corpus. The humans, who were exposed to my collected text material, are my readers. And while they are reading the text material, I can record their eye movements with an eye-tracker. In psycholinguistics, this method has been used for quite some time now. However, up until the last ten years or so, only carefully manipulated experimental stimuli were shown to readers. In psycholinguistic experiments, sentences might be constructed from scratch and only a single word could be exchanged for another to realise an experimental manipulation. Then, the effect of the experimental variation on processing behaviour is measured in form of reading variables (see §2.2 for a short introduction of reading variables).

If I record eye movements on natural text (i.e. the corpus I constructed), I can add those eye movements as an additional annotation layer. Now, a specific word is not only associated with a certain frequency measure, part-of-speech information and the phrase it is located in. It now also has processing information associated with it. With such a “reading corpus”, I have a powerful instrument at hand to investigate human reading of natural texts. The linguistic information can still be used to select specific instances from the corpus, but now I also know how humans processed these instances when they read them in the context of the whole corpus. Collection of eye-tracking data is expensive (in terms of scientific staff, participants and time). That is the reason why reading corpora are a lot smaller than text corpora we are used to in linguistics.

Several reading corpora are available in the field of psycholinguistics. Most of them are in English (Frank et al. 2013; Kennedy 2003), but there are also reading corpora for French (Kennedy 2003) and German (Kliegl et al. 2004; Wolfer et al. 2013). In this article, I present analyses based on the Freiburg Legalese Reading Corpus (FLRC), a corpus of jurisdictional terminological language. Jurisdictional language is known (at least in Germany) for its difficulties on several linguistic levels. Hansen-Schirra & Neumann (2004: 170) following Oksaar (1988) and Wagner (1981) identify the following linguistic properties as representative for jurisdictional language: long sentences, personalisations of inanimate objects or circumstances, complexity induced by derivations (the creation of new words with affixes), chains of subordinated nouns, extensive genitive attribution, archaic forms, formulaic expressions and nominalisations instead of verbs. I will focus on the last-mentioned structures: nominalisations. The overarching research

questions will be: Are nominalisations really harder to process than other words (especially than nouns, which are most comparable)? And can nominalisations be reformulated effectively to make text processing less complex?

I will start out by presenting some arguments why jurisdictional language should also be understandable to lay people (§1.1). I will then describe some reading corpora in more detail (§1.2). Chapter 2 will introduce the Freiburg Legalese Reading Corpus and the linguistic information (§2.1) and eye-tracking data it contains (§2.2). In Chapter 3, I will present my data selection and preparation processes (§3.1), formulate the hypotheses (§3.2) and present the statistical analyses and results (§3.3). I will discuss these results in Chapter 4. Chapter 5 will conclude the article.

1.1 Optimising the comprehensibility of jurisdictional language

The linguistic inaccessibility of jurisdictional language stands in contrast to the highly relevant function the jurisdictional system fulfils in modern democracies. Jurisdictional texts of all kinds ensure peaceful coexistence in our society. In Germany, the Bundesverfassungsgericht (Federal Constitutional Court) plays a prominent role in the German jurisdictional system and in society as a whole. On a linguistic level, its decisions are highly complex and not easily understandable for their ultimate addressees, the citizens of Germany who are mainly lay people when it comes to interpreting jurisdictional texts (cf. Eichhoff-Cyrus, Antos & Schulz 2009). Because of this prominent role of the Bundesverfassungsgericht, excerpts from decisions by this court were included into the reading corpus I am going to describe in this article. The corpus also contains full-length decisions. However, this part of the corpus will be of secondary interest in the present article.

Of course, one could say that jurisdictional language does not have to be understandable to lay people because it is a language for special purposes or professional jargon just like the language of, for example, IT staff, miners or linguists. Towfigh (2009) formulates such a position. There are (at least) two arguments against such a position. The professional jargons of IT people, miners or linguists are by far not as socially relevant as jurisdictional language. Of course, information technology also gets more important in modern social life. But still, it is far from being as important for the organisation of our social coexistence as jurisdictional language. If IT jargon might eventually get equally important, also the terminology of IT experts would face the demand of the public to get more understandable to everyone.

The second argument against the position that Towfigh (2009), amongst others, formulates is that even experts of the jurisdictional system struggle with their own professional code (cf. Eichhoff-Cyrus & Strobel 2009). As expected, 70% of all 84 surveyed legal experts judge their professional code as “weniger gut verständlich” (not well comprehensible) for lay people. Another 26% judge it as “nicht verständlich” (not comprehensible at all) for lay people (Eichhoff-Cyrus & Strobel 2009: 138). More surprisingly, though, 73% of the legal experts have *sometimes* trouble comprehending jurisdictional language themselves. Another 12% state that they *often* have trouble understanding jurisdictional language (Eichhoff-Cyrus & Strobel 2009: 146). The experts also largely agree that jurisdictional texts should be comprehensible for lay people without special training. Only 6% of all experts do not agree to this statement (Eichhoff-Cyrus & Strobel 2009: 139).

So, the demand for easier comprehensible jurisdictional language is formulated both by lay people and legal experts. With this motivation in mind, the Freiburg Legalese Reading Corpus was compiled. The goal was to provide detailed empirical data on the comprehension process of jurisdictional language. In terms of internal/external validity, some compromises were made in the Freiburg Legalese Reading Corpus: The part of the corpus with reformulations is clearly more similar to a classic psycholinguistic experiment because excerpts from original court decisions were reformulated by linguists making at least the reformulations not ecologically valid anymore. The other part of the corpus contains complete texts that were not altered. So, this part of the corpus can be considered ecologically more valid.

With this empirical data, hypotheses regarding the processing of specific linguistic constructions can be tested. Linguistic stimuli were not constructed from scratch but real-life linguistic stimuli were used, the aim being to reach high ecological validity. Hopefully, insights gained from these analyses can thus be generalised to other real-life texts.

1.2 Reading corpora

As already outlined in the introduction, research using reading corpora has gained increasing influence in psycholinguistics and related disciplines. Reading corpora are large collections of eye-tracking data on text material. There are already several reading corpora available in the field: The English UCL Corpus (Frank et al. 2013) contains eye-tracking and self-paced reading data. The German Potsdam Sentence Corpus (PSC, Kliegl, Nuthmann & Engbert 2006) consists of artificial sentences constructed around target words. These target words were selected for

word class (noun or verb), frequency (high or low) and length (short, medium or long). For each of the 12 combinations of these factors, 12 sentences were constructed, leading to a total of 144 sentences. Please keep in mind that these sentences never occurred in the real world and are not connected to each other on a content level. The English/French Dundee Corpus (Kennedy 2003) contains editorials from *The Independent* and *Le Monde*. Recently, the PopSci corpus with German popular science texts has been introduced to the field (Müller-Feldmeth, Wolfer & Konieczny 2013). Here, 16 texts from popular science journals are contained in the text corpus and were read by human readers. So, connections between sentences on a textual level (e.g. co-reference chains) were still intact.

One of the first applications of reading corpora has been the evaluation of models of eye-movement control. There is still a considerable debate in this field, mainly between two models, the E-Z Reader (Reichle, Rayner & Pollatsek 2003) and the SWIFT (Engbert et al. 2005) model. Fairly low-level processes are the focus of both computational models. The models are mainly interested in when and where a reader moves the eyes while reading text. These low-level processes have to be considered, of course, but are not the primary focus of this chapter. Recently, higher-level processes in language comprehension have come to the attention of researchers using reading corpora. Several psycholinguistic models and theories have been investigated and evaluated using reading corpora: surprisal (Demberg & Keller 2008; Patil, Vasishth & Kliegl 2009), cue-based parsing and similarity-based interference (Müller-Feldmeth, Wolfer & Konieczny 2013), semantic constraint (Pynte, New & Kennedy 2008) and many more. In this article, I am going to investigate research questions dealing with the lexical level. Namely, I am going to analyse the processing of nominalisations and how, if at all, they can be reformulated.

2 The Freiburg Legalese Reading Corpus

All data for the Freiburg Legalese Reading Corpus was collected in the eye-tracking labs of the Centre for Cognitive Science at the University of Freiburg.

2.1 Language material

The Freiburg Legalese Reading Corpus consists of two main parts: (1) A sub-corpus with nine original full length texts (three decisions, three press releases, and three newspaper articles) and (2) a sub-corpus containing thirty short sections of original decisions with thirty moderately reformulated texts and thirty

Table 1: Corpus design of the corpus part with the original excerpts and reformulations. Total token counts of the 10 texts in each cell are also shown. Column names are the three complexity types Hansen et al. (2006) used to select the original excerpts.

	Nominalisations	Complex NPs	Syntax
Original excerpts	10 texts 200 tokens	10 texts 303 tokens	10 texts 434 tokens
Moderate reformulations	10 texts 214 tokens	10 texts 317 tokens	10 texts 439 tokens
Strong reformulations	10 texts 217 tokens	10 texts 334 tokens	10 texts 440 tokens

strongly reformulated texts. The first part consisting of full-length texts has 12,769 tokens. The second part with excerpts from decisions and the associated reformulations has 2,898 tokens. In the analyses, I will use data from both corpus parts but will focus on the second part (excerpts and reformulations) later on.

The excerpts and reformulations are especially useful for the analysis at hand. See Table 1 for a brief overview over the design of this corpus part. The thirty original excerpts were selected to meet one of three types of linguistic complexity (henceforth “complexity type”). Ten texts contain many nominalisations that were transformed into verbal structures during the course of reformulation (complexity type “nominalisations”). Ten texts contain very complex noun phrases – mostly due to excessive pre- or post-nominal modification (complexity type “complex NPs”). In the first reformulation step, these modifications were transformed into subordinate clauses. In the second step leading to the strongly reformulated version, sentences were split to avoid the sentential complexity induced by these subordinate sentences. The remaining ten original texts contained very long and multiply embedded sentences (complexity type “syntax”). They were split up repeatedly to achieve the moderately and strongly reformulated versions. All texts, including the reformulated versions, were taken from a study of Hansen et al. (2006). They reformulated the texts with the help of a jurisdictional expert who made sure that semantic content of the texts was maintained. Annotations of the texts were added by two student assistant annotators with the help of the software “Annotate” (Plaehn 1998) for semi-automatic syntactic annotation.

Parts-of-speech according to the Stuttgart-Tübingen-TagSet¹ and phrase structure were annotated. “Annotate” makes suggestions for both and the annotators altered these annotations in case they were incorrect. When the annotators did not agree on a certain annotation, agreement was established through discussion. Nominalisations were identified manually.

Example 1 shows an original excerpt from the complexity type “nominalisations” (1a) and the associated strongly reformulated text (1b)². From Example 1a, the original excerpt, in transition to Example 1b, the strong reformulation, Hansen et al. (2006) transformed four *ung-nominalisations (*Herstellung, Reduzierung, Aufhebung, Aufhebung*) into verbal structures (*hergestellt wurde, reduziert/aufgehoben wird, aufzuheben*).

- (1) Mehr als zwölf Jahre nach der Herstellung der deutschen
More than twelve years after the realisation the.GEN German.GEN
Einheit habe sich die Ermächtigung zur Reduzierung oder
unity.GEN have.SBJV itself the authorization to reduction or
Aufhebung der Gebührenermäßigung durch Rechtsverordnung
abolishment the.GEN reduction.of.fees through legal.decree
zu einer Rechtspflicht zur Aufhebung des
to a legal.duty for.the abolishment the.GEN
Gebührenabschlags verdichtet.
partial.payment.of.fees condensed.
‘More than twelve years after the realisation of German reunification, the
authorization for reduction or abolishment of a rebate of fees by decree is
said to have been condensed to a legal duty of abolishment of rebates.’
- (2) Mehr als zwölf Jahre nachdem die deutsche Einheit hergestellt wurde,
More than twelve years after the German unity realised was,
habe sich die Ermächtigung, dass die Gebührenermäßigung
have.SBJV itself the authorization, that the reduction.of.fees
durch Rechtsverordnung reduziert oder aufgehoben wird, zu der
through legal.decree reduced or abolished is.AUX, to the

¹ The tag table is available under <http://www.ims.uni-stuttgart.de/forschung/ressourcen/lexika/TagSets/stts-table.html> (Accessed 2014-12-23)

² The English meaning of this excerpt is: More than twelve years after Germany was re-united, the authorisation to reduce or repeal the reduction of fees by legal decree has been condensed to a legal duty to repeal the partial payment of fees.

Rechtspflicht verdichtet, den Gebührenabschlag aufzuheben.
legal.duty condensed the partial.payment.of.fees abolish.

‘More than twelve years after German reunification was realised, the authorization to reduce or abolish a rebate by decree is said to have been condensed to a legal duty of abolishing rebates.’

As already mentioned above, this does not change the semantic content of the text. Example 1 is a prototypical example in terms of reformulations of nominalisations. The frequently used nominalisations are transformed into verbal structures. However, as can also be seen in Example 1, subordinating structures have to be introduced to integrate these newly introduced verbal structures in the sentence context. These subordinating structures can be identified by the subordinating conjunctions *nachdem* and *dass* in Example 1b. Only nouns ending with *-ung* are treated as nominalisations in the remainder of this article.

When I analyse all 30 texts within the complexity type nominalisations (first column in Table 1), I see that the share of nominalisations in all words drops significantly from 16.8 % (originals) to 10.6 % (moderate reformulations) and 10.5 % (strong reformulations). Raw numbers for all 10 texts in each cell taken together are 35 (originals), 23 (moderate reformulations) and 23 (strong reformulations) nominalisations. So, obviously, no nominalisations were reformulated during the second step of the reformulation process³. To measure the impact of reformulations on the use of verbal structures, I sum up all occurrences of verbs and participles and look at the development of shares and raw figures over the reformulation versions. The verbal structures follow the opposite pattern of nominalisations. The share of verbal structures in all words rises significantly from 9.7 % (originals) to 20.0 % (moderate reformulations) and 21.1 % (strong reformulations). Again, most of the reformulations were obviously made between the originals and the moderate reformulations. The raw figures confirm this. The original texts contain 18, the moderately reformulated texts 41 and the strongly reformulated versions 45 verbs and/or participles. In the remainder of this chapter, I will investigate which influences these reformulations have on the reading process.

2.2 Eye-tracking data

All texts were distributed on pages that matched a 20-inch-flatscreen with a resolution of 1600 by 1200 pixels. Texts were presented in a 48 pt proportional serif

³ For some of the texts, Hansen et al. (2006) did not create a strongly reformulated version.

font to allow for as natural as possible a reading experience. A maximum of 11 lines of text with 1.5 line spacing fit on one screen page. After each text, a comprehension question had to be answered by pressing ‘yes’ or ‘no’ on a response box. Questions were rather easy and were primarily included into the study to keep participants attentive. I used an SR Research EyeLink 1000 for data collection. The eye-tracker measured gaze position of the participants with a rate of 1000 Hz which produces 1 data point every millisecond. Calibration and validation was carried out before the experiment and – if necessary – data collection was interrupted to recalibrate the eye-tracker.

Reading data was collected from 80 human readers (40 for each corpus part) who were given course credit or monetary compensation for their participation. Participants were all students at the University of Freiburg with normal (56 participants) or corrected vision (24 participants). Data on the age of participants was not gathered. It was made sure that none of the participants had an educational background in law. 51 participants were female. Participants were seated with their heads on a chin rest, so that their eyes were approximately 60 cm away from the screen.

Reading time variables were calculated with custom R (R Core Team 2014) scripts from the fixation data collected and pre-calculated by the eye-tracker’s data export tool (SR Research DataViewer). For the identification of fixations, I used the default parameter settings of DataViewer. Each fixation was associated to an interest area. Interest areas spanned individual words and were expanded vertically to the middle of the space between lines. Fixations not associated with an interest area were discarded.

The following reading variables can be considered standard measures in psycholinguistic reading research and were pre-computed for all words in the corpus: first fixation duration (the duration in milliseconds of the first fixation on the current word), first-pass reading times (the summed duration of all fixations from entering the word until exiting the word to the left or to the right), regression path durations (the summed duration of all fixations from entering the word until a word right of the word is fixated, including all fixations on material left of the word, i.e. regressions) and total reading times (the summed duration of all fixations on the word, also including fixations after the word has been exited to the right for the first time, i.e. if the reader re-reads the word later on). Several more reading variables can be calculated from these measures. For example, re-reading time (total reading time minus first-pass reading time) or a binary variable if a regressive saccade has been launched during the first reading of a word (is the regression path duration longer than the first-pass reading time?). Another vari-

able that is frequently used is skipping probability that is derived from a binary variable if a word is read during first pass (is the first-pass reading time greater than zero?). Considering the whole FLRC, mean first fixation duration was 200 msec. Mean first-pass reading time was 265 msec. Mean regression path durations were 495 msec. Mean total reading times were 405 msec. On average, a regressive saccade was launched on 12.6% of all words in the FLRC.

I will not go into much detail regarding the cognitive processes associated with each of these reading time variables. This has been excellently done elsewhere in much more detail (Clifton, Staub & Rayner 2007; Hyönä, Lorch & Rinck 2003; Rayner & Pollatsek 2006). A few words have to be said, though. First-pass reading times include all fixations during reading a region of interest (here, regions of interest are words) for the first time. First-pass reading times therefore capture processes of word recognition and early stages of word processing. Regression path duration or go-past time, as it is also called, “can reasonably be construed as the time it takes upon reading the target word on first pass until it is successfully integrated with the ongoing context” (Rayner & Pollatsek 2006: 620). This, however, should not be the main problem of processing nominalisations. Of course, nominalisations have to be integrated into the context of the sentence – just as any other noun. However, I expect that the specific problem of processing a nominalisation has to do with derivation, and not with integration into the previous sentence context. So, I expect effects of nominalisations in rather early measures (e.g. first-pass reading time) and not in regression path durations. If I indeed find an effect in regression path durations and the probability of launching a regressive saccade upon encountering a nominalisation, this may be a hint that nominalisations are indeed harder to integrate into the previous context than normal nouns. An effect in total reading times cannot be ruled out because nominalisation may also be revisited after first pass.

3 The impact of nominalisations on the reading process

Nominalisations can be considered complex for two reasons. Hansen et al. (2006) describe them as an instrument to increase the informational density in a text: “Information is packed into heavy noun phrases and nominalisations rather than being distributed onto larger grammatical units [...]” (ibid., p. 24). Also, nominalisations can be used to create constructions like “die Abschiebung wird durchgeführt” (the deportation is carried out). Such constructions enable objective and pertinent descriptions (“objektive und sachbezogene Darstellung“, Hansen-Schirra & Neumann 2004: 169) without mentioning a specific agent of the action. This

may be odd for lay people and may lead to comprehension difficulties because such agent-less constructions are not very common in every-day language but rather have to be considered a property often found in terminological language (cf. Hansen-Schirra & Neumann 2004: 170).

Given the previously introduced linguistic material and the corresponding processing data, I am able to investigate several questions. The most obvious question is: Is it difficult to process nominalisations? This, however, is not a valid research question because no linguistic entity is mentioned to which nominalisations are compared. I will start by comparing nominalisations to all other nouns. I will then also compare the size of the effect nominalisations have on the reading process with the effects of other word classes (content/function words, finite verbs, all nouns). With this comparison, I want to estimate the relative size of the effect that nominalisations have on the reading process.

I will then go into detail using the part of the reading corpus that contains excerpts and reformulations. Ten texts were specifically selected because they contain many nominalisations. Those texts were reformulated in two steps. I will therefore check if the reformulations generally slow down the processing of verbs and participles. The final questions will be if overall text processing benefits from the reformulations of nominalisations. All research questions are also summarised in the first column of Table 2.

3.1 Data selection and preparation

Each research question posed above needs a specific subset of the reading corpus to be answered. The data subsets associated with the respective research questions can be found in the second column of Table 2.

After the respective data subset has been selected from the reading corpus, I have to control for some effects that are rather obvious but not interesting for the research questions at hand. For example, it is common sense that longer words take longer to read. Also, it has been shown repeatedly that corpus frequency is a good predictor for reading times. The more frequent the actual word (hitherto: word n) is encountered in natural language, the faster it is read (cf. Kliegl et al. 2004). This frequency effect extends to bigram and trigram frequencies, the corpus frequency of word n in combination with words $n - 1$ and, for trigrams, $n - 2$ (cf. Boston et al. 2008). Another factor that plays a crucial role for reading times is orthographic familiarity. Orthographic familiarity is a measure of how often a specific word shape appears in the language (cf. White 2008). Orthographic familiarity is operationalised by the cumulative frequency of all words with the same initial three letters and the same length as the critical word.

Table 2: Research questions, selected data subsets, and hypotheses for the analyses presented in this chapter

Research question	Data subset	Hypothesis
Are nominalisations harder to process than normal nouns?	All nouns (including nominalisations) in the FLRC	Nominalisations are read longer during first-pass.
How pronounced is the nominalisation effect compared to other classes of words?	All words in the FLRC	Compared to other word classes, the nominalisation effect is rather large.
Are nouns read faster after nominalisations have been reformulated?	All nouns in complexity type nominalisations	Nouns are read faster after nominalisations have been transferred to verbal structures.
Do reformulations shift complexity to verbal structures?	Verbs and participles in complexity type nominalisations	No observable effect
Does overall text processing benefit from the reformulations?	Whole texts of complexity type nominalisations	Reformulated texts are processed faster than original excerpts.

On the syntactic level, the position of word n in the sentence and its depth of embedding in the phrase structure are relevant predictors for reading times (cf. Pynte, New & Kennedy 2008). Words that occur later in the sentence are read faster. The same holds for words that are embedded deeper in the syntactic structure.

Other factors that can influence the reading times of words are related to the presentation of the words on the screen. The position of the text on the screen is relevant because, generally speaking, words appearing “later” on the screen are read faster. Also, reading behaviour on first and last words in lines can deviate from standard reading behaviour within a line. This is especially true for the first word in a row because it is the first word that is encountered after a change of

text lines. If a reader changes the line, she or he has to initiate a very long saccade from the end of the line to the beginning of the next line. The fixation following this long saccade tends to be longer.

I control for all these factors by including the relevant variables word length, unigram, bigram and trigram frequencies, orthographic familiarity, position in the sentence, depth of embedding and presentation factors in statistical baseline models⁴. Participant and text identity were included as random intercepts into the models. When I analyse the effect of nominalisation, for example, the effect of the word's frequency is controlled for beforehand. This is desirable because, on the one hand, it reduces noise in my data and could lead to more pronounced effects of the factors I am interested in. On the other hand, this procedure also makes sure that variance in reading times that is really explained by word frequency or another control factor is not erroneously ascribed to other effects. Intuitively, the procedure leads to “cleaner” data and more reliable models.

3.2 Hypotheses

All hypotheses can also be seen in the third column of Table 2. If nominalisations are indeed more complex than normal nouns, I expect elevated first-pass (and eventually total) reading times for nominalisations compared to normal nouns. I would not necessarily expect an effect in regression path durations because the main problem of nominalisations should be a lexical one and not to integrate the nominalisations into the sentence context. At least, this should not be more of a problem than for normal nouns. When I compare the effect of nominalisations to the effects of other word classes, I would expect that the nominalisation effect is rather large. One could say that, the larger the nominalisation effect is in comparison to other effects, the more important it is to reformulate nominalisations or not to introduce them into texts in the first place. I will evaluate these first two hypotheses in the first section of the results section.

Regarding the reformulations, I would expect nouns to be read faster after nominalisations have been reformulated. This should simply be the case because the share of nominalisations in all nouns gets lower in reformulated texts (see §2.1 for statistics on the linguistic consequences of reformulations). The effect should be observable for first-pass (and eventually) total reading times.

I would also expect – if the reformulations were successful – that complexity is not simply shifted to another linguistic level or towards other linguistic

⁴ The statistical baseline models are quite extensive and can be requested from the author. All models in the current article were fitted using linear mixed-effects models within the statistical environment R (R Core Team 2014) and the package lme4 (Bates et al. 2014).

structures, namely verbs and participles (see §2.1 for the relationship between reformulations and verbs and participles). If reformulations do not slow down processing of verbal structures, I do not expect an effect of reformulations on the reading times of verbs and participles. I will evaluate the latter three hypotheses in the second section of the results section.

For the last research question, I move away from the level of single words and take the whole texts of the complexity type nominalisations into account. This is a rather coarse grained measure that cannot tell us anything about the processing of single words. However, I can investigate the overall time it needs to process the original and reformulated texts as a whole. If reformulations were successful also on the textual level, I would expect slightly reduced reading times for the whole text for the reformulated versions.

3.3 Results

In this section, the research questions and associated hypotheses are categorised into two groups. The first group of research questions relates to the complexity of nominalisations in the whole reading corpus and the comparison to other word classes. The second group of research questions deals with the consequences it has on text processing when nominalisations are reformulated.

3.3.1 Complexity of nominalisations

For the first research questions, all nouns in the reading corpus were selected. Each of those nouns is associated with the information if it is a nominalisation or not. This is the only predictor I include into my model as a fixed effect. I calculated linear mixed-effects models for the reading time variables first-pass reading time, total reading time and regression path durations. Those reading time variables were corrected by the baseline model procedure introduced in §3.1. For skipping probability and the probability that a regressive saccade is launched upon encountering the noun, logistic mixed regression models were calculated. These models are better suited for binary outcome variables (cf. Jaeger 2008).

3 shows the model parameters for this first analysis. First-pass reading times, total reading times and also regression path durations are significantly higher for nouns that are nominalisations than for normal nouns. Please bear in mind that the effects for the reading time variables cannot simply be ascribed to word length (nominalisations are likely to be longer than normal nouns) because the baseline modelling rules this out beforehand. The probability of being skipped is significantly lower for nominalisations. The probability that a regressive saccade

Table 3: Model parameters (estimate, standard error and t value) for the fixed effect of a noun being a nominalisation. The five models for the different reading variables are in rows. P values are only available for the logistic mixed regression models for skipping probability and the probability of a regressive saccade (last two rows). For all other models, an absolute t value above 2 strongly indicates significance. The sign of the estimate indicates the effect direction.

	Estimate	Standard Error	t/z value
First-pass reading times	0.063	0.012	5.243
Total reading times	0.131	0.015	8.911
Regression path durations	0.073	0.020	3.615
Probability of being skipped	-0.536	0.033	-16.087
			$p < .0001$
Probability that a regressive saccade is launched	0.059	0.035	1.662
			$p = 0.10$

is launched upon encountering a nominalisation is *not* significantly higher than for normal nouns. This is somewhat surprising because the regression path durations are significantly higher for nominalisations. A likely explanation would be that first-pass reading times are also included in regression path durations and that the effect in regression path durations are really ascribable to the effect in first-pass reading times. So, I can conclude that nominalisations do not trigger more regressions back into previous text material than normal nouns. This relates to the question if nominalisations may also be harder to integrate into the sentence context than normal nouns. Given the results in Table 3, this does not seem to be the case. Nevertheless, the impact of nominalisations on context integration processes needs to be further investigated to tease apart these effects.

With the next analysis, I am comparing the nominalisation effect to other sets of words. I will concentrate on total reading times here, because the most pronounced effect was shown for this variable (second row of Table 3). I am going to compare the effect estimates because this model parameter can be thought of as an operationalisation of the effect strength⁵. I include all words in the reading corpus into this analysis. For each word, I code if it belongs to one or more of the following set of words: content words, function words, finite verbs, nouns, nominalisations. I chose finite verbs as a set for comparison because finite verbs

⁵ In simple linear regression with only one predictor variable, the estimate is the slope of the regression line.

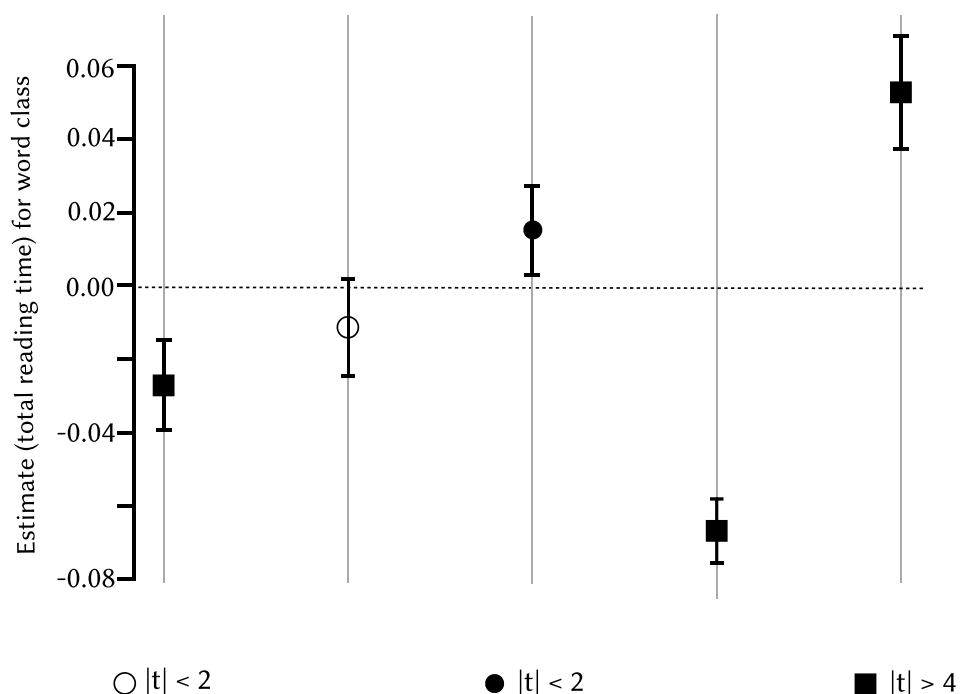


Figure 1: Comparison of effects estimates (y axis) for total reading times and the effects of being a content word, function word, finite verb, noun or nominalisation. The associated t value is symbolized by the plot symbols.

most of the time encode the main action going on in a sentence and should be considered quite relevant. With this coding, it is clear that each nominalisation belongs not only to the set of nominalisations but are also a subset of content words and nouns.

All those declarations that are either TRUE or FALSE for each word then enter one linear mixed-effects model as fixed effects. By comparing the estimates for the fixed effects, I get an impression of the relative importance of these estimates. The estimates are visualised in Figure 1. We are especially interested in the right-most point, the effect estimate for nominalisations. The estimate ($\beta = 0.053$, $t = 6.80$) is relatively high compared to all other effect estimates. This means that total reading times on a word are much higher when it is a nominalisation. This is surprising when I compare this estimate with the one for all nouns. The relatively low estimate ($\beta = -0.066$, $t = -15.15$) for nouns means that, if a word is a noun, total reading times on this word are much lower than for all other words. It can also be seen that the effect for nominalisations seems to be way stronger than

the one for finite verbs ($\beta = 0.016$, $t = 2.55$) which have to be considered quite important elements in sentences.

3.3.2 Consequences of reformulating nominalisations

The results for the analyses of reformulations do not seem to be as clear as the ones introduced in the last chapter. If the reading times for the baseline models are used, no significant effects can be shown. All differences go into the right direction here (nouns are read faster in reformulated versions) but there seems to be too little explanatory power in the data to show significant effects here. If only word length, word frequency, word familiarity and presentation factors are included (see §3.1 for an explanation of these variables), a significant effect of reformulation can be shown. Since most of the reformulation is going on between the original excerpts and the moderately reformulated versions, both moderately and strongly reformulated versions are treated as “reformulated”. This way, I only have the contrast between nouns in original and nouns in reformulated texts. This makes the model parameters interpretable more easily. The model shows that nouns in reformulated texts are indeed read faster. This time, I only find a significant effect for first-pass reading times ($\beta = -0.076$, $t = -2.05$) but not for total reading times ($\beta = -0.080$, $t = -1.22$). The next question is if processing complexity shifts to verbs and participles when nominalisations are reformulated. For this analysis, I selected all verbs and participles⁶ from the texts of complexity type “nominalisations”. Verbs and participles were selected because nominalisations were transformed into verbal structures (see Example 1). It would be unfortunate for reformulation efforts if processing complexity simply shifts to the “substitute structures” that are introduced when nominalisations are reformulated (also see Table 2). No effects can be shown for either first-pass reading times ($\beta = -0.011$, $t = -0.20$) or total reading times ($\beta = -0.106$, $t = -1.65$) and all effect estimates point into the negative direction, i.e. verbs and participles are read slightly (and not significantly) *faster*. Please note, if no effect can be shown, this does not mean that there really is no effect. It could always be the case that I simply cannot *detect* the effect. However, if reformulations really would lead to increased reading times on verbs and participles because these were the structures that were introduced by transforming nominalisations, then the reading times in the reformulated version should be *higher*, which is not the case.

⁶ In the tag set I used, verbs and participles are all words from one of the following parts-of-speech: VAFIN, VMFIN, VVINFIN, VVFIN, VVIZU, VAINF, VVPP, VMPP, VAPP.

To answer the last question, I take the analyses to the text level and compare the reading times of the different reformulation versions of the texts from the complexity type nominalisations. Indeed, as Table 4 shows, the mean reading time per text is getting lower as texts are reformulated. However, the noise (standard deviations, second column) is also very high here. Apart from that, I also have to take text length into account, because it is not that big of a surprise that texts are read faster if they are shorter. Indeed, they get a little shorter in average as the third column of Table 4 shows. To tease apart the effect of text

Table 4: Mean text reading times, standard deviations and mean text lengths for texts with complexity type nominalisations

	Mean text reading time	Standard deviation	Mean text length
Original excerpts	9489 ms	7607 ms	149 characters
Moderate reformulations	8487 ms	7748 ms	144 characters
Strong reformulations	8018 ms	6193 ms	143 characters

length and relate the noise to the effect, I have to calculate a statistical model. Reformulation version and text length (as a control factor) are entered as fixed effects and participants and text identity are treated as random intercepts. I predict the logarithm of text reading time⁷. As expected, there is a clear effect of text length on text reading time ($\beta = 0.006$, $t = 5.16$), texts are read longer when they contain more characters. Apart from that, the effect I am interested in is also significant and it points into the direction the hypothesis suggests. Strongly reformulated texts are read faster than the original excerpts ($\beta = -0.157$, $t = -2.58$). Moderately reformulated texts, however, lie somewhere in between the originals and the strong reformulations ($\beta = -0.109$, $t = -1.79$) and are not significantly different from either one of them in terms of text reading time. Again, this does not replicate the findings of Hansen et al. (2006). There are several possible reasons for this: (1) I used a different method to measure text reading time. I summed all total reading times of all words. Hansen et al. (2006) measured the time between text presentation on-set till participants advanced to the question. (2) Hansen et al. did not correct for text length when estimating reading times. (3) Also, to my knowledge, they did not include random intercepts for participants and/or text

⁷ The effects I report are even more pronounced when raw text reading time is predicted. However, it is statistically more appropriate to predict logarithmized text reading times because they follow a normal distribution more closely than raw text reading times.

identity. That means that inter-individual differences (between texts and/or participants) are not accounted for. (4) Hansen and colleagues analyse reading time differences over all complexity types while I only analysed texts of the complexity type “nominalisations” because these are the texts relevant for the current research question.

4 Discussion

All hypotheses mentioned in the third column of Table 2 were confirmed by the statistical analyses. Indeed, nouns, which are nominalisations, are read slower. This is confirmed by significant effects for first-pass reading times and total reading times. Also, nominalisations are less likely to be skipped. Although nouns in general are read faster than all other words in the reading corpus, this effect is the other way around for nominalisations. The effect size for nominalisation (i.e. the impact on the reading process) is comparable and even higher than the same effect for finite verbs.

So, nominalisations really seem to be quite complex for the reader to process. However, I also showed that – at least in my reformulation corpus – they can be reformulated. In the reformulated versions of the original excerpts, nominalisations were transformed into verbal structures. When comparing the reading times of nouns in the original versions with those in the reformulated versions, I found significantly faster reading on all nouns. It also does not seem to be some kind of trade-off where I just switch complexity introduced by nominalisations for complexity on verbs and participles. Those did not seem to be read longer in reformulated text versions. When using the very rough measure of overall text reading time, I saw that reformulated texts are read slightly faster. So, although nominalisations indeed seem to introduce a fair bit of complexity into jurisdictional texts, they obviously can be reformulated without just shifting processing complexity to another linguistic level.

However, I still have to bring these results into line with the larger context of text processing and text comprehension. It has been shown during the course of this article that nominalisations take longer to process and this is a quite pronounced effect. However, I did not present any data on the consequences for the mental representation the participants built up while reading the texts. Such consequences could be measured by asking participants to answer questions after reading the texts. This also has been done during data collection for the FLRC. I used the same questions like Hansen et al. (2006). However, the questions did not seem to be sensitive enough to measure improvements (or declines) in comprehension performance. Questions after all versions of the texts were answered

similarly well. 84% of all questions after original excerpts were answered correctly, which is already a quite good result. For reformulated versions, this figure only rose marginally to 88% for moderately reformulated texts and 87% correctly answered questions for strongly reformulated texts. This does not replicate the results of Hansen et al. (2006) who found overall significantly better results for the moderate reformulations (approximately⁸ 85% correct answers) than for the original excerpts (75%) and the strongly reformulated versions (78%). I do not have a clear explanation for these differences.

Note, however, that Hansen et al. only report overall results, i.e. the other complexity types with complex noun phrases and complex syntax are also included in their analysis. As a consequence, it is not possible to compare my results concerning only the complexity type “nominalisations” with the results of Hansen et al. (2006).

5 Conclusion

Jurisdictional texts are complex on many linguistic levels. One of the main difficulties is a large amount of nominalisations. Indeed, those seem to be associated with slower comprehension processes. Fortunately, I also showed that this complexity could be resolved by transforming the nominalisations into verbal structures. However, more research is necessary to also investigate the consequences for the mental representations readers construct in their minds while reading jurisdictional texts. Just because some parts of the texts are read considerably slower does not mean that they are not comprehended at all. For this, better tests measuring the products of comprehension processes have to be developed. With the empirical data presented in this article, I have to assume that optimised texts (and also portions of text) are read faster – if this is also associated with a better understanding of text content remains to be shown.

References

Bates, Douglas, Martin Maechler, Ben Bolker, Steven Walker, Rune Haubo Bojesen Christensen & Henrik Singmann. 2014. *Lme4: Linear mixed-effects models using eigen and s4* [software manual]. R package version 1.1-6. <http://cran.r-project.org/web/packages/lme4/index.html>.

⁸ Figures have to be extracted from the plot from Hansen et al. (2006: 35). Raw figures are not available in the text.

- Boston, Marisa, John Hale, Reinhold Kliegl, Umesh Patil & Shravan Vasishth. 2008. Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *The Mind Research Repository (beta)* 2(1). 1–12.
- Clifton, Charles, Adrian Staub & Keith Rayner. 2007. Eye movements in reading words and sentences. In Roger P. G. van Gompel, Martin H. Fischer, Wayne S. Murray & Robin L. Hill (eds.), *Eye movements: A window on mind and brain*, 341–371. Amsterdam: Elsevier.
- Demberg, Vera & Frank Keller. 2008. Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition* 109(2). 193–210.
- Eichhoff-Cyrus, Karin M., Gerd Antos & Rüdiger Schulz. 2009. *Wie denken die Deutschen über die Rechts- und Verwaltungssprache? Eine repräsentative Umfrage der Gesellschaft für deutsche Sprache*. Wiesbaden: Gesellschaft für deutsche Sprache.
- Eichhoff-Cyrus, Karin M. & Thomas Strobel. 2009. Einstellungen der Justiz zu Rechts- und Verwaltungssprache. Eine Trendumfrage. *Der Sprachdienst* 53(5). 133–151.
- Engbert, Ralf, Antje Nuthmann, Eike M. Richter & Reinhold Kliegl. 2005. SWIFT: A dynamical model of saccade generation during reading. *Psychological Review* 112(4). 777–813.
- Frank, Stefan L., Irene Fernandez Monsalve, Robin L. Thompson & Gabriella Vigliocco. 2013. Reading time data for evaluating broad-coverage models of English sentence processing. *Behavior Research Methods* 45(4). 1182–1190.
- Hansen, Sandra, Ralph Dirksen, Martin Kuchler, Kerstin Kunz & Stella Neumann. 2006. Comprehensible legal texts-utopia or a question of wording? On processing rephrased German court decisions. *Hermes* 36. 15–40.
- Hansen-Schirra, Silvia & Stella Neumann. 2004. Linguistische Verständlichmachung in der juristischen Realität. In Kent D. Lerch (ed.), *Recht verstehen. Verständlichkeit, Missverständlichkeit und Unverständlichkeit von Recht*, 167–184. Berlin: de Gruyter.
- Hyönä, Jukka, Robert F. Lorch & Mike Rinck. 2003. Eye movement measures to study global text processing. *The mind's eye: Cognitive and applied aspects of eye movement research*. Jukka Hyönä, Ralph Radach & Heiner Deubel (eds.). 313–334.
- Jaeger, T. Florian. 2008. Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language* 59(4). 434–446.
- Kennedy, Alan. 2003. The dundee corpus. [CD-ROM]. School of Psychology, The University of Dundee.

- Kliegl, Reinhold, Antje Nuthmann & Ralf Engbert. 2006. Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General* 135(1). 12–35.
- Kliegl, Reinhold, Ellen Grabner, Martin Rolfs & Ralf Engbert. 2004. Length, frequency, and predictability effects of words on eye movements in reading. *European Journal of Cognitive Psychology* 16(1-2). 262–284.
- Müller-Feldmeth, Daniel, Sascha Wolfer & Lars Konieczny. 2013. The influence of pre-verbal referents on verb processing: Evidence for similarity-based interference from a new German reading corpus. In C. Frenck-Mestre, F.-X. Alario, N. Nguyen, P. Blache & C. Meunier (eds.), *Proceedings of the 19th Conference on Architectures and Mechanisms for Language Processing (AMLaP)*, 263–287. Berlin: Frank & Timme.
- Oksaar, Els. 1988. *Fachsprachliche Dimensionen*. Vol. 4. Tübingen: Gunter Narr.
- Patil, Umesh, Shravan Vasishth & Reinhold Kliegl. 2009. Compound effect of probabilistic disambiguation and memory retrievals on sentence processing: Evidence from an eyetracking corpus. In Andrew Howes, David Peebles & Richard Cooper (eds.), *Proceedings of 9th international conference on cognitive modeling*, 92–97. <http://eprints.hud.ac.uk/10530/>, accessed 2017-07-26.
- Plaehn, Oliver. 1998. Annotate Programm-Dokumentation (negra project report). Saarbrücken. <http://www.coli.uni-saarland.de/projects/sfb378/negra-corpus/annotate-manual.ps.gz>, accessed 2015-10-21.
- Pynte, Joel, Boris New & Alan Kennedy. 2008. A multiple regression analysis of syntactic and semantic influences in reading normal text. *Journal of Eye Movement Research* 2(1). 1–11.
- R Core Team. 2014. *A language and environment for statistical computing [software manual]*. <http://www.r-project.org>, accessed 2016-07-18.
- Rayner, Keith & Alexander Pollatsek. 2006. Eye movement control in reading. In Matthew J. Traxler & Morton A. Gernsbacher (eds.), *Handbook of Psycholinguistics*. 2nd edn. 613–657. Amsterdam: Elsevier.
- Reichle, Erik D., Keith Rayner & Alexander Pollatsek. 2003. The EZ reader model of eye-movement control in reading: Comparisons to other models. *Behavioral and Brain Sciences* 26(4). 477–526.
- Towfigh, Emanuel Vahid. 2009. Komplexität und Normenklarheit – oder: Gesetze sind für Juristen gemacht. *Der Staat* 1(48). 29–73.
- Wagner, Hildegard. 1981. *Die deutsche Verwaltungssprache der Gegenwart: Eine Untersuchung der sprachlichen Sonderform und ihrer Leistung*. 3rd edn. (Schriften des Instituts für deutsche Sprache 9). Düsseldorf: Schwann.

- White, Sarah J. 2008. Eye movement control during reading: Effects of word frequency and orthographic familiarity. *Journal of Experimental Psychology: Human Perception and Performance* 34(1). 205–223.
- Wolfer, Sascha, Daniel Müller-Feldmeth, Lars Konieczny, Uli Held, Karin Maksymski, Silvia Hansen-Schirra, Sandra Hansen & Peter Auer. 2013. PopSci: A reading corpus of popular science texts with rich multi-level annotations. A case study. In *Book of abstracts of the 17th European Conference on Eye Movements*. Lund.

